SHORT REPORT
doi: 10.2306/scienceasia1513-1874.2006.32.421

ScienceAsia 32 (2006): 421-427

Identification of the Polyhedrin Gene of Thai *Bombyx mori* Nucleopolyhedrovirus

Mallika Kaewwises[1,2], Sudawan Chaeychomsri[3] , Srimek Chowpongpang[4] and Tipvadee Attathom[5*]

[1] Center for Agricultural Biotechnology, Kasetsart University, Kamphaeng Saen, Nakhon Pathom 73140, Thailand.
[2] Biotechnology Research and Development Office, Sirindhon Plant Genetic Resources Building, Thanyaburi, Pathum Thani 12110, Thailand.
[3] Central Laboratory and Greenhouse Complex, Kasetsart University, Kamphaeng Saen, Nakhon Pathom 73140, Thailand.
[4] Department of Plant Pathology, Faculty of Agriculture, Kasetsart University, Chatuchak, Bangkok 10900, Thailand.
[5] Department of Entomology, Faculty of Agriculture Kamphaeng Saen, Kasetsart University, Kamphaeng Saen, Nakhon Pathom 73140, Thailand.

* Corresponding author, E-mail: agrtva@ku.ac.th

Received 19 Aug 2005
Accepted 5 Jun 2006

ABSTRACT: The full-length of the polyhedrin gene (*polh*) of Thai *Bombyx mori* nucleopolyhedrovirus (*Bm*NPV) was cloned and sequenced. The *polh* sequence contained a 735 bp open reading frame (ORF) encoding a protein of 245 amino acids with a predicted molecular mass of 28.8 kDa. The nucleotide sequence of Thai *Bm*NPV *polh* shows greater than 98% identity to the sequences of five different *Bm*NPV *polh* genes that were previously characterized. The high degree of sequence identity with the *polh* sequences of other *Bm*NPVs suggests that they are orthologues of the *Bm*NPV *polh* gene in this study. Comparison of Thai *Bm*NPV *polh* sequence with other *polh*s of Lepidoteran NPVs (*Autographa californica*, *Helicoverpa armigera*, *Spodoptera litura* and *S. exigua*) indicated that the nucleotide and amino acid sequences were greater than 65% and 78% identical, respectively.

KEYWORDS: nucleopolyhedrovirus, polyhedrin gene, *Bombyx mori*, silkworm, Thai.

INTRODUCTION

Bombyx mori nucleopolyhedrovirus (*Bm*NPV) is an infectious agent causing the most destructive disease (grasserie) of the silkworm, *B. mori*. *Bm*NPV belongs to the genus Nucleopolyhedrovirus, the family Baculoviridae[1]. Baculoviruses have been detected in over 600 species of arthropod hosts including members of the order Lepidoptera, Diptera, Hymenoptera, Coleoptera, Neuroptera, Thysanura and Trichoptera[2]. The two morphological subgroups within the NPVs are the single-embedded NPV, in which only one nucleocapsid is present per envelope, and the multiple-embedded NPV, in which several nucleocapsids are packed per envelope[3]. Polyhedra (occlusion bodies) produced during an epizootic outbreak may persist between seasons in the environment and therefore serve as a reservoir of inoculum to infect subsequent generations of insect hosts[4]. Polyhedra are mainly composed of a single polypeptide known as polyhedrin (Polh). Polh, which constitutes the crystalline matrix of baculovirus occlusion bodies, plays a significant role in the replication cycle of baculovirus[5]. Polh is encoded by a gene which is highly conserved among baculoviruses, so it is the most comprehensive option available for estimating the relationship among baculoviruses[6].

More than 80% sequence identity has been reported among lepidopteran baculovirus Polhs[7]. Based on phylogenetic studies using the amino acid sequences of polyhedron proteins, lepidopteran NPVs have been classified into two groups, namely Group I and Group II[6]. Subsequently, Bulach *et al.*[8] supported these clades and revealed other subclades within Group II by analyzing the polyhedrin and DNA polymerase genes. Evolutionary studies of baculoviruses have been carried out with several additional genes, but systematics of these viruses is still consistent.

The *polh* genes of NPVs have been characterized in many host insects, such as *A. californica, Anticarsia gemmatalis, B. mori, H. armigera*, *Lymantria dispar, Orgyia pseudotsugata, S. exigua, S. litura*, *S. frugiperda*, *S. littoralis*, etc[9]. However, the *polh* gene of the Thai *Bm*NPV isolate has not been investigated. This paper aimed to study

polh from the Thai *Bm*NPV isolate and determine its relatedness to other *Bm*NPVs. Information on the relationships among *Bm*NPV *polh* and other NPV *polhs* is necessary for the understanding of NPV evolution. Moreover, the information is useful for study of the molecular basis of NPVs and their useful applications in construction of recombinant baculoviruses with higher insecticidal activity and for detection of *Bm*NPV for preventive control of grasserie disease.

Materials and Methods

Purification of *Bm*NPV and DNA Extraction

*Bm*NPV that was isolated from diseased larvae collected from Udon Thani Sericultural Research Center, Thailand was propagated in silkworm larvae. The polyhedra were purified by 40-65% (w/w) sucrose density gradient centrifugation[10]. The virions were released from the polyhedra by dissolving with alkaline solution (0.2 M Na_2CO_3, 0.5 M EDTA, 0.34 M NaCl). Genomic DNA was extracted according to the method of O'Reilly *et al*.[11].

PCR Amplification and Cloning of *Bm*NPV *polh*

A primer set (F: 5'- CCCAAGATGTATAAACCA- 3' and R: 5'- GCCTAACGCGCCCGATGTT- 3') was designed from the nucleotide sequence of *Bm*NPV T3 (GenBank accession number L33180)[12]. The PCR amplification cycling was an initial denaturation at 94 °C for 5 min, followed by 35 repeated cycles of 94 °C for 1 min, 56 °C for 1 min, 72 °C for 1 min, and the final extension at 72 °C for 5 min in a PCT-100 DNA thermal cycler (MJ Research, Inc.). The PCR product was ligated into pGEM®-T vector (Promega) and cloned. The plasmid was purified using standard techniques[13], and subjected to sequencing.

Sequence Comparison and Phylogenetic Analysis of the *polh* Genes from Different NPVs

Nucleotide sequences of the coding region and the predicted amino acid sequences of *polh* were analyzed using Lasergene®(DNASTAR, Inc.). Thai *Bm*NPV *polh* was compared with five other *Bm*NPV *polh genes*: Japanese 1 (T3; Acession number L33180)[12], Japanese 2 (M30925)[14], Korean (U75359), Chinese (X63614) and Canadian (M100430)[15]. In addition, *polh* nucleotide and amino acid sequences of other lepidopteran NPVs namely: *A. californica* NPV (*Ac*NPV; NC_001623)[16], *H. armigera* NPV (*Ha*NPV; NC_002654)[17], *S. exigua* NPV (*Se*NPV; NC_002169)[18] and *S. litura* NPV (*Sl*NPV; NC_003102)[19] were included in the analysis. A phylogenetic tree was constructed using the neighbor joining method[20]. To determine the reproductivity values for the clustering within the tree, the original data was bootstrapped 1,000 times.

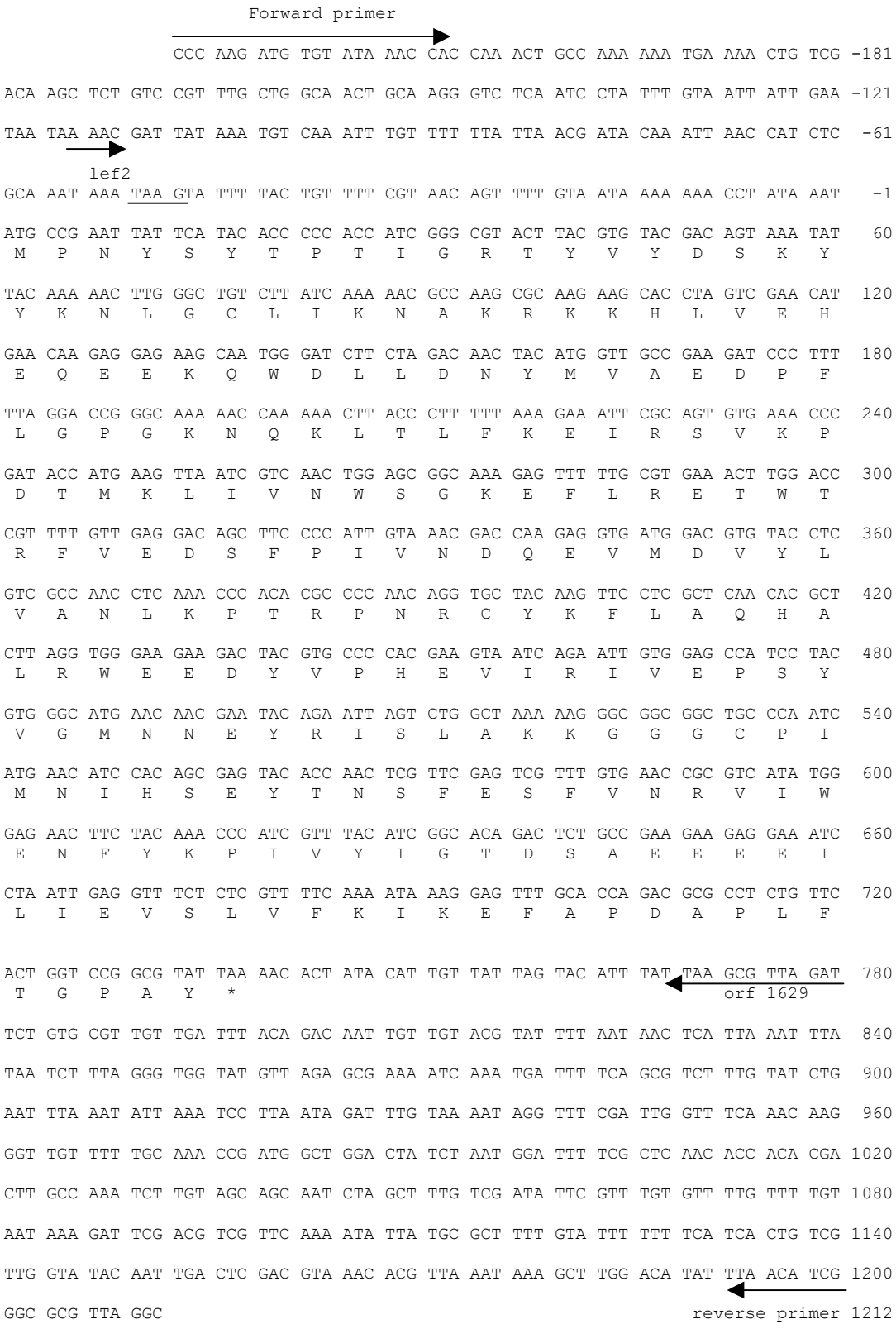
Results and Discussion

Nucleotide Sequence of Thai *Bm*NPV *polh*

A 1,440-bp fragment of the Thai *Bm*NPV DNA that contained the full-length coding region of the *polh* (GenBank accession number AY779044) was successfully amplified and sequenced (Fig 1). The *polh* ORF consisted of 735 nucleotides that encoded a polypeptide of 245 amino acids with the predicted molecular mass of 28.8 kDa. The sequence contained 228 bp of 5' UTR and 474 bp of 3' UTR, respectively. Several characteristics of the Thai *Bm*NPV *polh* gene sequence were investigated. The immediate upstream sequence of the translation initiation site was AT rich and contained the uniquely conserved transcription start site TAAG motif, which is similar to the consensus sequence of TAAATAAGTATTTT at position -42 to -56 of other baculovirus late gene promoters[21]. DNA sequences similar to the consensus TATA and CAAT, which represent important elements of eukaryotic gene promoters were observed at positions -107 and -143. The canonical poly (A) signal AATAAA was present in the 3'end of the Thai *Bm*NPV *polh* gene at position 1081. The Thai *Bm*NPV *polh* ORF had the translation initiation codon, ATG, and the termination codon, TAA, as found in other baculoviruses[11]. The complete sequence of the Thai *Bm*NPV *polh* overlapped with the *lef*-2 (late gene expression factor-2) gene in the 5' flanking region and with *orf*1629 in the 3' flanking region. The *lef*-2 gene was located in the *polh* upstream region, adjacent to the transcription start site in the same direction and the *orf*1629 was located in the *polh* downstream region in the reverse direction.

The orientation of *polh*, *lef*-2 and *orf*1629 in *Bm*NPV are the same as those found in the *Ac*NPV genome[16]. The position of the *orf1629* gene in *Ha*NPV, *Se*NPV and *Sl*NPV is next to *polh*, similar to *Bm*NPV, but *lef*-2 is not located in the *polh* upstream region. In *Bm*NPV, *lef*-2 is essential for both viral DNA replication and late gene expression[22] and *orf*1629 is essential for *Bm*NPV viability[23].

The flanking nucleotides of Thai *Bm*NPV *polh* and the *Bm*NPV T3 *polh* fragment (accession number L33180), which was used to design the primers, were compared. The major difference between the two isolates was the deletion of eight nucleotides in the upstream region of Thai *Bm*NPV *polh* ORF at the position -71 and 3 substitutions at positions A^{-119}/G^{-101}, C^{-128}/T^{-110}, C^{-154}/T^{-136} were found, while the rest of the nucleotide sequences were almost identical. A previous study using a series of deletions in the upstream region of *Bm*NPV *polh* (*lef*-2, *orf*327, *orf*453 and *bro-e*) revealed that the upstream region of *polh* has no effect on expression from the *polh* promoter[24].




```
                    CCC AAG ATG TGT ATA AAC CAC CAA ACT GCC AAA AAA TGA AAA CTG TCG -181

ACA AGC TCT GTC CGT TTG CTG GCA ACT GCA AGG GTC TCA ATC CTA TTT GTA ATT ATT GAA -121

TAA TAA AAC GAT TAT AAA TGT CAA ATT TGT TTT TTA TTA ACG ATA CAA ATT AAC CAT CTC  -61

GCA AAT AAA TAA GTA TTT TAC TGT TTT CGT AAC AGT TTT GTA ATA AAA AAA CCT ATA AAT   -1

ATG CCG AAT TAT TCA TAC ACC CCC ACC ATC GGG CGT ACT TAC GTG TAC GAC AGT AAA TAT   60
 M   P   N   Y   S   Y   T   P   T   I   G   R   T   Y   V   Y   D   S   K   Y

TAC AAA AAC TTG GGC TGT CTT ATC AAA AAC GCC AAG CGC AAG AAG CAC CTA GTC GAA CAT  120
 Y   K   N   L   G   C   L   I   K   N   A   K   R   K   K   H   L   V   E   H

GAA CAA GAG GAG AAG CAA TGG GAT CTT CTA GAC AAC TAC ATG GTT GCC GAA GAT CCC TTT  180
 E   Q   E   E   K   Q   W   D   L   L   D   N   Y   M   V   A   E   D   P   F

TTA GGA CCG GGC AAA AAC CAA AAA CTT ACC CTT TTT AAA GAA ATT CGC AGT GTG AAA CCC  240
 L   G   P   G   K   N   Q   K   L   T   L   F   K   E   I   R   S   V   K   P

GAT ACC ATG AAG TTA ATC GTC AAC TGG AGC GGC AAA GAG TTT TTG CGT GAA ACT TGG ACC  300
 D   T   M   K   L   I   V   N   W   S   G   K   E   F   L   R   E   T   W   T

CGT TTT GTT GAG GAC AGC TTC CCC ATT GTA AAC GAC CAA GAG GTG ATG GAC GTG TAC CTC  360
 R   F   V   E   D   S   F   P   I   V   N   D   Q   E   V   M   D   V   Y   L

GTC GCC AAC CTC AAA CCC ACA CGC CCC AAC AGG TGC TAC AAG TTC CTC GCT CAA CAC GCT  420
 V   A   N   L   K   P   T   R   P   N   R   C   Y   K   F   L   A   Q   H   A

CTT AGG TGG GAA GAA GAC TAC GTG CCC CAC GAA GTA ATC AGA ATT GTG GAG CCA TCC TAC  480
 L   R   W   E   E   D   Y   V   P   H   E   V   I   R   I   V   E   P   S   Y

GTG GGC ATG AAC AAC GAA TAC AGA ATT AGT CTG GCT AAA AAG GGC GGC GGC TGC CCA ATC  540
 V   G   M   N   N   E   Y   R   I   S   L   A   K   K   G   G   G   C   P   I

ATG AAC ATC CAC AGC GAG TAC ACC AAC TCG TTC GAG TCG TTT GTG AAC CGC GTC ATA TGG  600
 M   N   I   H   S   E   Y   T   N   S   F   E   S   F   V   N   R   V   I   W

GAG AAC TTC TAC AAA CCC ATC GTT TAC ATC GGC ACA GAC TCT GCC GAA GAA GAG GAA ATC  660
 E   N   F   Y   K   P   I   V   Y   I   G   T   D   S   A   E   E   E   E   I

CTA ATT GAG GTT TCT CTC GTT TTC AAA ATA AAG GAG TTT GCA CCA GAC GCG CCT CTG TTC  720
 L   I   E   V   S   L   V   F   K   I   K   E   F   A   P   D   A   P   L   F


ACT GGT CCG GCG TAT TAA AAC ACT ATA CAT TGT TAT TAG TAC ATT TAT TAA GCG TTA GAT  780
 T   G   P   A   Y   *                                               orf 1629

TCT GTG CGT TGT TGA TTT ACA GAC AAT TGT TGT ACG TAT TTT AAT AAC TCA TTA AAT TTA  840

TAA TCT TTA GGG TGG TAT GTT AGA GCG AAA ATC AAA TGA TTT TCA GCG TCT TTG TAT CTG  900

AAT TTA AAT ATT AAA TCC TTA ATA GAT TTG TAA AAT AGG TTT CGA TTG GTT TCA AAC AAG  960

GGT TGT TTT TGC AAA CCG ATG GCT GGA CTA TCT AAT GGA TTT TCG CTC AAC ACC ACA CGA 1020

CTT GCC AAA TCT TGT AGC AGC AAT CTA GCT TTG TCG ATA TTC GTT TGT GTT TTG TTT TGT 1080

AAT AAA GAT TCG ACG TCG TTC AAA ATA TTA TGC GCT TTT GTA TTT TTT TCA TCA CTG TCG 1140

TTG GTA TAC AAT TGA CTC GAC GTA AAC ACG TTA AAT AAA GCT TGG ACA TAT TTA ACA TCG 1200

GGC GCG TTA GGC                                                 reverse primer 1212
```

Fig 1. Nucleotide sequence of Thai *Bm*NPV polyhedrin gene and its flanking regions (1,440 bp, AY 779044). The predicted amino acid sequence is indicated by one- letter code. The sequence at nt -51 initiates the 5' end of the mRNA. The putative transcription initiation motif (TAAG) is underlined. The PCR primers are also shown and the arrows indicate the direction of extension.

Comparison of Thai *Bm*NPV *polh* with Other NPV *polh* genes

All *Bm*NPV *polh* ORFs described in this study contained 735 nucleotides that encoded polypeptides of 245 amino acids. The ORFs of the *Bm*NPV *polh* of the Thai and Japanese 1 (T3) isolates were identical. Nucleotide and amino acid sequences of Thai *Bm*NPV *polh* revealed high identity (more than 95%) to those of other *Bm*NPVs and lower identity with *Ac*NPV, *Ha*NPV, *Se*NPV and *Sl*NPV (75.2%, 65.3%, 70.1% and 65.9% for *polh* gene sequences and 86.1%, 78.0%, 81.6% and 79.6% for Polh protein sequences, respectively; Table 1). This result is similar to a previous report, which indicated that the percent identities of the amino acid sequence of the Canadian *Bm*NPV Polh compared to those of *Ac*NPV, *Se*NPV, and *Sl*NPV were 86%, 82% and 80%, respectively[25].

Alignment of the *Bm*NPV Polh sequences showed variation of amino acid sequences near the N-terminus (Fig 2). Comparison of the sequence divergence of nucleotide and amino acid sequences among *Bm*NPVs indicated that *Bm*NPVs found in the Asian countries (Thailand, Japan, Korea and China) were more closely related than the isolate from Canada.

The ORF of *Bm*NPV *polh* encoded a polyhedrin protein containing 245 amino acids, similar to *polh* of *Ac*NPV and *Ha*NPV, while the *polh* ORF of *Se*NPV and *Sl*NPV contained 246 and 249 amino acids, respectively. The amino acid sequence alignment of *Bm*NPV Polh with the other NPVs demonstrated that more differences occur in the N-terminus than the C-terminus (Fig 2). In *Ha*NPV, the amino acid (Histidine/Aspartic acid) of Polh at position 40 was not found. There are many substitutions that make amino acids of *Bm*NPVs differ from other NPVs such as V^{31}/L, L^{51}/W, M^{107}/V, F^{123}/Y, M^{128}/L, R^{129}/K, C^{147}/W, D^{148}/E, P^{149}/E, L^{187}/I, I^{199}/V and L^{226}/I. Interestingly, the amino acid sequences at position 147-149, of all *Bm*NPV studied were different from other NPVs. This sequence may be used to differentiate *Bm*NPVs and NPVs from other host species. There is no report on mutation of amino acids at positions 147-149 of polyhedrin. Therefore, the functionally importance of these amino acid domains is not known. Amino acid content of the putative *Bm*NPV Polh indicated that it is rich in acidic amino acid residues, such as glutamic acid, especially at position 220-223 where four glutamic acids occur consecutively.

The phylogenetic tree of NPV Polh proteins showed that NPV could be divided into two groups (Fig 3). The first group was further divided to 2 subgroups; *Bm*NPV-Canada (I.1) and the Asian *Bm*NPV (I.2) whereas the second group was composed of *Sl*NPV, *Ha*NPV, *Se*NPV and *Ac*NPV. Among the *Bm*NPVs, the Canadian isolate is more distantly related than the Asian isolates (Fig 3). Due to limited vagility of the silkworm, biogeographic variation was observed between different isolates of NPVs in this species, as demonstrated by phylogeographic differences between the Canadian and Asian isolates.

Nucleopolyhedrovirus clades based on the Polh were first described by Zanotto *et al*. [6], who divided NPVs into Group I and Group II. Subsequently, Bulach *et al*.[8] supported these clades and revealed other subclades within Group II by analyzing the combined Polh and DNA polymerase sequences. Our results were concordant with Zanotto *et al*.[6] and Bulach *et al*.[8] where *Ha*NPV, *Se*NPV and *Sl*NPV were allocated into Group II. However, *Bm*NPVs revealed possibly paraphyletic relationships between the Canadian and the Asian isolates.

Table 1. Nucleotide (above) and deduced amino acid sequence (below diagonal) identity of *Bm*NPVs *polh*. Sequence data of *Bm*NPV *polh* varieties included 1) Thai (this study, AY779044); 2) Japanese 1 (T3) (L33180); 3) Japanese 2 (M30925); 4) Korean (K1) (U75359); 5) Chinese (X63614); 6) Canadian (M100430); 7) *Ac*NPV (NC_001623); 8) *Ha*NPV (NC_002654); 9) *Se*NPV (NC_002169) and 10) *Sl*NPV (NC_003102).

Percent nucleotide sequence identity											
		1	2	3	4	5	6	7	8	9	10
***Bm*NPV-Thai**	1		100	99.9	99.7	99.6	98.4	75.2	65.3	70.1	65.9
***Bm*NPV-Japanese 1 (T3)**	2	100		99.9	99.7	99.6	98.4	75.2	65.3	70.1	65.9
***Bm*NPV-Japanese 2**	3	99.6	99.6		99.6	99.5	98.5	75.0	65.0	69.8	65.6
***Bm*NPV-Korean**	4	99.6	99.6	99.2		99.3	98.1	75.2	65.3	69.9	65.7
***Bm*NPV-Chinese**	5	99.6	99.6	99.2	99.2		98.0	74.8	65.2	69.7	65.4
***Bm*NPV-Canadian**	6	96.7	96.7	97.1	96.3	96.3		74.6	64.6	70.5	65.9
***Ac*NPV**	7	86.1	86.1	85.7	85.7	85.1	86.1		67.1	72.5	72.0
***Ha*NPV**	8	78.0	78.0	77.6	77.6	77.6	77.1	81.6		70.6	70.1
***Se*NPV**	9	81.6	81.6	81.2	81.2	81.2	81.6	84.9	84.1		76.0
***Sl*NPV**	10	79.6	79.6	79.2	79.6	79.2	79.6	84.1	82.9	84.1	
Percent amino acid sequence identity											

```
BmNPV-Thai                M----PNYSYTPTIGRTYVYDNKYYKNLGCLIKNAKRKKHLVEHEQEEKQ     50
BmNPV-Japanese 1 (T3)     M----PNYSYTPTIGRTYVYDNKYYKNLGCLIKNAKRKKHLVEHEQEEKQ     50
BmNPV-Japanese 2          M----PNYSYTPTIGRTYVYDNKYYKNLGCLIKNAKRKKHLVEHEQEEKQ     50
BmNPV-Korean (K1)         M----PNYPYTPTIGRTYVYDNKYYKNLGCLIKNAKRKKHLVEHEQEEKQ     50
BmNPV-Chinese             M----PNYSYTPTIGRTYVYDNKYYKNLGCLIKNAKRKKHLVEHEQEEKQ     50
BmNPV-Canadian            M----PNYSYNPTIGRTYVYDNKYYKNLGGLIKNAKRKKHLIEHEKEEKQ     50
AcNPV                     M----PDYSYRPTIGRTYVYDNKYYKNLGAVIKNAKRKKHFAEHEIEEAT     50
HaNPV                     M---YTRYSYSPTLGKTYVYDNKYFKNLGAVIKMPTQEH-LEEHEHEERN     50
SeNPV                     M---YTRYSYNPALGRTYVYDNKFYKNLGSVIKNAKRKEHLLQHEIEERT     50
SlNPV                     MYSRYSAYNYSPHLGKTYVYDNKYYKNLGHVIKNAKRKHDALEREADERE     50
Consensus                 M    PNYSYTPTIGRTYVYDNKYYKNLGCLIKNAKRKKHLVEHEQEEKQ     50

BmNPV-Thai                WDLLDNYMVAEDPFLGPGKNQKLTLFKEIRSVKPDTMKLIVNWSGKEFLR    100
BmNPV-Japanese 1 (T3)     WDLLDNYMVAEDPFLGPGKNQKLTLFKEIRSVKPDTMKLIVNWSGKEFLR    100
BmNPV-Japanese 2          WDLLDNYMVAEDPFLGPGKNQKLTLFKEIRSVKPDTMKLIVNWSGKEFLR    100
BmNPV-Korean (K1)         WDLLDNYMVAEDPFLGPGKNQKLTLFKEIRSVKPDTMKLIVNWSGKEFLR    100
BmNPV-Chinese             WDLLDNYMVAQDPFLGPGKNQKLTLFKEIRSVKPDTMKLIVNWSGKEFLR    100
BmNPV-Canadian            WDLLDNYMVAEDPFLGPGKNQKLTLFKEVRNVKPDTMKLIVNWSGKEFLR    100
AcNPV                     LDPLDNYLVAEDPFLGPGKNQKLTLFKEIRNVKPDTMKLVVGWKGKEFYR    100
HaNPV                     LDSLDKYLVAEDPFLGPGKNQKLTLFKEIRSVKPDTMKLVVNWSGREFLR    100
SeNPV                     LDPLERYVVAEDPFLGPGKNQKLTLFKEIRIVKPDTMKLVVNWSGKEFLR    100
SlNPV                     LDHLDKYLVAEDPFMGPGKNQKLTLFKEIRNVKPDTMKLIVNWNGKEFLR    100
Consensus                 WDLLDNYMVAEDPFLGPGKNQKLTLFKEIRSVKPDTMKLIVNWSGKEFLR    100

BmNPV-Thai                ETWTRFVEDSFPIVNDQEVMDVYLVANLKPTRPNRCYKFLAQHALRWEED    150
BmNPV-Japanese 1 (T3)     ETWTRFVEDSFPIVNDQEVMDVYLVANLKPTRPNRCYKFLAQHALRWEED    150
BmNPV-Japanese 2          ETWTRFVEDSFPIVNDQEVMDVYLVANLKPTRPNRCYKFLAQHALRWEED    150
BmNPV-Korean (K1)         ETWTRFVEDSFPIVNDQEVMDVYLVANLKPTRPNRCYKFLAQHALRWEED    150
BmNPV-Chinese             ETWTRFVEDSFPIVNDQEVMDVYLVANLKPTRPNRCYKFLAQHALRWEED    150
BmNPV-Canadian            ETWTRFVEDSFPIVNDQEVMDVYLVANLKPTRPNRCYKFLAQHALRWDED    150
AcNPV                     ETWTRFMEDSFPIVNDQEVMDVFLVVNMRPTRPNRCYKFLAQHALRCDPD    150
HaNPV                     ETWTRFMEDSFPIVNDQEIMDVFLSVNMRPTKPNRCYRFLAQHALRCDPD    150
SeNPV                     ETWTRFMEDSFPIVNDQEIMDVFLVINMRPTRPNRCFRFLAQHALRCDPD    150
SlNPV                     ETWTRFMEDSFPIVNDQEVMDVFLVVNMRPTRPNRCFRFLAQHALRCDPE    150
Consensus                 ETWTRFVEDSFPIVNDQEVMDVYLVANLKPTRPNRCYKFLAQHALRWEED    150

BmNPV-Thai                YVPHEVIRIVEPSYVGMNNEYRISLAKKGGGCPIMNIHSEYTNSFESFVN    200
BmNPV-Japanese 1 (T3)     YVPHEVIRIVEPSYVGMNNEYRISLAKKGGGCPIMNIHSEYTNSFESFVN    200
BmNPV-Japanese 2          YVPHEVIRIMEPSYVGMNNEYRISLAKKGGGCPIMNIHSEYTNSFESFVN    200
BmNPV-Korean (K1)         YVPHEVIRIVEPSYVGMNNEYRISLAKKGGGCPIMNIHSEYTNSFESFVN    200
BmNPV-Chinese             YVPHEVIRIVEPSYVGMNNEYRISLAKKGGGCPIMNIHSEYTNSFESFVN    200
BmNPV-Canadian            YVPHEVIRIMEPSYVGMNNEYRISLAKKGGGCPIMNIHSEYTNSFESFVN    200
AcNPV                     YVPHDVIRIVEPSWVGSNNEYRISLAKKGGGCPIMNLHSEYTNSFEQFID    200
HaNPV                     YIPHEVIRIVEPSYVGSNNEYRISLAKKYGGCPVMNLHAEYTNSFEDFIT    200
SeNPV                     YVPHEVIRIVEPVYVGTNNEYRISLAKKGGGCPVMNLHSEYTNSFEEFIN    200
SlNPV                     YVPHDVIRIVEPSYVGTNNEYRISLAKKGGGCPVMNLHAEYTTSFESFID    200
Consensus                 YVPHEVIRIVEPSYVGMNNEYRISLAKKGGGCPIMNIHSEYTNSFESFVN    200

BmNPV-Thai                RVIWENFYKPIVYIGTDSAEEEEILIEVSLVFKIKEFAPDAPLFTGPAY     249
BmNPV-Japanese 1 (T3)     RVIWENFYKPIVYIGTDSAEEEEILIEVSLVFKIKEFAPDAPLFTGPAY     249
BmNPV-Japanese 2          RVIWENFYKPIVYIGTDSAEEEEILIEVSLVFKIKEFAPDAPLFTGPAY     249
BmNPV-Korean (K1)         RVIWENFYKPIVYIGTDSAEEEEILIEVSLVFKIKEFAPDAPLFTGPAY     249
BmNPV-Chinese             RVIWENFYKPIVYIGTDSAEEEEILIEVSLVFKIKEFAPDAPLFTGPAY     249
BmNPV-Canadian            RVIWENFYKPIVYIGTDSAEEEEILIEVSLVFKIKEFAPDAPLFTGPAY     249
AcNPV                     RVIWENFYKPIVYIGTDSAEEEEILLEVSLVFKVKEFAPDAPLFTGPAY     249
HaNPV                     NVIWENFYKPIVYVGTDSAEEEEILLEVSLIFKIKEFAPDAPLYTGPAY     249
SeNPV                     RVIWENFYKPIVYVGTDSGEEEEILLELSLVFKIKEFAPDAPLYNGPAY     249
SlNPV                     KVIWYNFYKPIVYVGTDSAEEEEILLEVSLVFKIKEFAPDAPLYTGPAY     249
Consensus                 RVIWENFYKPIVYIGTDSAEEEEILIEVSLVFKIKEFAPDAPLFTGPAY     249
```

Fig 2. Multiple alignments of the deduced amino acids of Thai *Bm*NPV *polh* (AY779044) and other *Bm*NPV *polhs*; Japanese 1 (T3) (L33180), Japanese 2 (M30925), Korean (K1)(U75359), Chinese (X63614), Canadian (M100430) and *Ac*NPV (NC_001623), *Ha*NPV (NC_002654), *Se*NPV (NC_002169) and *Sl*NPV (NC_003102). Positions containing different amino acids are shaded.

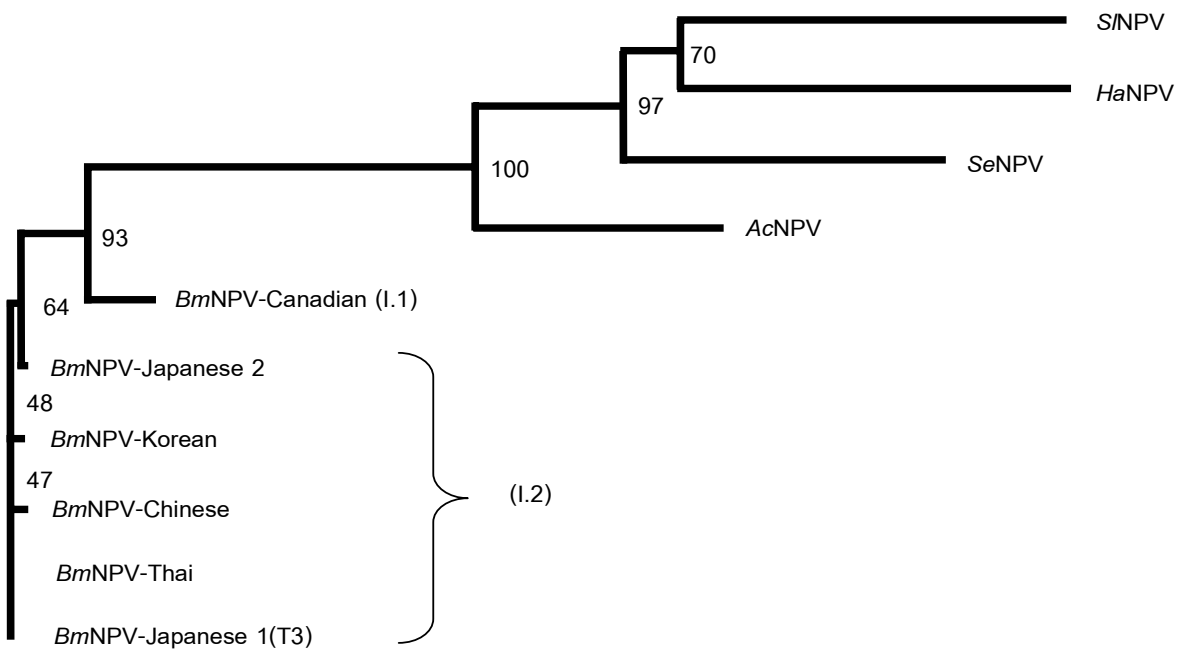


Fig 3. A phylogenetic tree constructed from deduced amino acid sequences of 6 *Bm*NPV *polhs* (see details from the legend of Fig 2), *Ac*NPV (NC_001623), *Ha*NPV (NC_002654), *Se*NPV (NC_002169) and *Sl*NPV (NC_003102). Numbers of the nodes indicate the percentage of reproductivity that resulted from 1000 bootstrap replicates of the original data.

Harrison and Bonning[26] constructed a phylogenetic tree of Polh of many lepidopteran NPVs including *Bm*NPV, *Ac*NPV, *Ha*NPV, *Se*NPV and *Sl*NPV. They grouped *Ha*NPV, *Se*NPV and *Sl*NPV into Group II and *Bm*NPV into Group I of the proposed tree of Zanotto *et al*.[6]. The *Ac*NPV polyhedrin was put on a branch outside of the clade containing the other members of Group I. They suggested that *Ac*NPV may have acquired its *polh* gene by recombination with another virus that is not closely related to other NPVs in Group I[26]. In addition, Jehle[27] used a Hidden Markov Model to surmise that the *Ac*NPV *polh* is a chimeric gene which consists of a mosaic of the genome of Group I and II NPVs. From these results, *Ac*NPV can be grouped in both Group I and Group II, depending on the method used for analysis. In this study, AcNPV was also placed in Group I. Since several reports revealed that adding the *Ac*NPV *polh* resulted in distortion and instability to the *polh* gene tree, many other genes were recently employed for phylogenetic analysis of baculovirus[28]. However, for simple molecular analysis, the *polh* is still useful because a great number of the *polh* gene sequences are available in the GenBank[29].

Acknowledgements

We would like to thank Busara Rawinoo from Department of Agriculture, Ministry of Agriculture and Cooperatives for providing silkworm eggs. This study was financially supported by the Center for Agricultural Biotechnology, Thailand.

References

1. van Regenmortel MHV and Fauquet C (2000) Virus taxonomy: classification and nomenclature of viruses: seventh report of the International Committee on Taxonomy of Viruses. Academic Press, San Diego, California.
2. Hong HK, Woo SD, Choi JY, Lee HK, Kim MH, Je YH and Kang SK (2000) Characterization of four isolates of *Bombyx mori* nucleopolyhedrovirus. *Arch Virol* **145,** 2351-61.
3. Hunter-Fujita FR, Entwistle PF, Evans HF and Crook NE (1998) Insect Viruses and Pest Management. Willey & Sons, Chichester.
4. Boucias DG, and Pendland JC (1998) Principles of insect pathology. Kluwer Academic Publishers.
5. Acharya A, Sriram S, Sehrawat S, Rahman M, Sehgal D and Gopinathan KP (2002) *Bombyx mori* nucleopolyhedrovirus: Molecular biology and biotecnological applications for large-scale synthesis of recombinant proteins. *Curr Sci* **28,** 455-65.
6. Zanotto PMA, Kessing BD and Maruniak JE (1993) Phylogenetic interrelationships among baculoviruses: Evolutionary rates and host associations. *J Invertebr Pathol* **62,** 147-64.
7. Rohrmann GF (1992) Baculovirus structure proteins. *J Gen Virol* **73,** 749-61.
8. Bulach DM, Kuman CA, Zaia A, Liang B and Tribe DE (1999) Group II nucleopolyhedrovirus subgroups revealed by phylogenetic analysis of polyhedrin and DNA polymerase

gene sequences. *J Invertebr Pathol* **73,** 59-73.
9. Woo SD (2001) Rapid detection of multiple nucleopolyhedroviruses using polymerase chain reaction. *Mol Cells* **11,** 334-40.
10. Attathom T, Chaeychomsri S, Chaichuchot S, Attathom S and Chiemsobat P (1988) Characterization of the nuclear polyhedrosis virus of the cotton bollworm, *Heliothis armigera. Kasetsart J (Nat Sci Suppl)* **22,** 14-23.
11. O'Reilly DR, Miller LK and Luckow VA (1992) Baculovirus Expression Vectors: Laboratory Manual. New York Oxford, Oxford University Press.
12. Gomi S, Majima K and Maeda S (1999) Sequence analysis of the genome of *Bombyx mori* nucleopolyhedrovirus. *J Gen Virol* **80,** 1323-37.
13. Sambrook J, and Russell DW (2001) Molecular Cloning: a laboratory manual. Cold Spring Habor, New York.
14. Maeda S, Kawai T, Obinata M, Horiuchi H, Saeki Y, Sato Y and Furusawa M (1985) Production of human alpha-interferon in silkworm using a baculovirus vector. *Nature* **315,** 592-4.
15. Iatrou K, Ito K and Witkiewicz H (1985) Polyhedrin gene of *Bombyx mori* nuclear polyhedrosis virus. *J Virol* **54,** 436-45.
16. Ayres MD, Howard SC, Kuzio J, Lopez-Ferber M, and Possee RD (1994) The complete DNA sequence of *Autographa californica* nuclear polyhedrosis virus. *Virology* **202,** 586-605.
17. Chen X, Ijkel WF, Tarchini R, Sun X, Sandbrink H, Wang H, Peters S *et al.* (2001) The sequence of the *Helicoverpa armigera* single nucleocapsid nucleopolyhedrovirus genome. *J Gen Virol* **82,** 241-57.
18. Ijkel WF, Strien EA, Heldens JG, Broer R, Zuidema D, Goldbach RW and Vlak JM (1999) Sequence and organization of the *Spodoptera exigua* multicapsid nucleopolyhedrovirus genome. *J Gen Virol* **80,** 3289-304.
19. Pang Y, Yu J, Wang L, Hu X, Bao W, Li G, Chen C *et al.* (2001) Sequence analysis of the *Spodoptera litura* multicapsid nucleopolyhedrovirus genome. *Virology* **287,** 391-404.
20. Saitou N, and Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Bol Evol* **4,** 406-25.
21. Leisy D, Rohrmann G and Beaudreau G (1986) The nucleotide sequence of the polyhedrin gene region from the multicapsid baculovirus of *Orgyia pseudotsugata*. *Virology* **153,** 280-8.
22. Sriram S and Gopinathan KP (1998) The potential role of a late gene expression factor, Lef2, from *Bombyx mori* nuclear polyhedrosis virus in very late gene transcription and DNA replication. *Virology* **251,** 108-22.
23. Je YH, Chang JH, Kim MH, Roh JY, Jin BR and O'Reilly DR (2001) The use of defective *Bombyx mori* nucleopolyhedrovirus genomes maintained in *Escherichia coli* for the rapid generation of occlusion-positive and occlusion-negative expression vectors. *Biotechnol Lett* **23,** 1809-17.
24. Acharya A and Gopinathan KP (2001) Identification of an enhancer-like element in the polyhedrin gene upstream region of *Bombyx mori* nucleopolyhedrovirus. *J Gen Virol* **82,** 2811-9.
25. Chou C, Huang C, Lo C, Kou G and Wang C (1996) Characterization of *Perida nuda* nucleopolyhedrovirus (*Penu*NPV) polyhedrin gene. *J Invertebr Pathol* **67,** 259-66.
26. Harrison RL and Bonning BC (2003) Comparative analysis of the genomes of *Rachiplusia ou* and *Autographa californica* multiple nucleopolyhedrovirus. *J Gen Virol* **84,** 1827-42.
27. Jehle JA (2004) The mosaic structure of the polyhedrin gene of the *Autographa californica* nucleopolyhedrovirus (*Ac*NPV). *Virus Genes* **29,** 5-8.
28. Herniou EA, Olszewski JA, Cory JS and O'Reilly DR (2004) The genome sequence and evolution of baculovirus. *Annu Rev Entomol* **48,** 211-34.
29. Lange M, Wang H, Zhihong H and Jehle JA (2004) Toward a molecular identification and classification system of lepidopteran-specific baculovirus. *Virology* **325,** 36-47.